Crowded, cell-like environment induces shape changes in aspherical protein

Dirar Homouz*†, Michael Perham†‡, Antonios Samiotakis*, Margaret S. Cheung*, and Pernilla Wittung-Stafshede‡§¶‖

*Department of Physics, University of Houston, Houston, TX 77204; and Departments of ‡Chemistry, §Biochemistry and Cell Biology, and ¶Keck Center for Structural Computational Biology, Rice University, Houston, TX 77251

Edited by Peter G. Wolynes, University of California at San Diego, La Jolla, CA, and approved July 1, 2008 (received for review April 15, 2008)

How the crowded environment inside cells affects the structures of proteins with aspherical shapes is a vital question because many proteins and protein–protein complexes *in vivo* adopt anisotropic shapes. Here we address this question by combining computational and experimental studies of a football-shaped protein (i.e., *Borrelia burgdorferi* VlsE) in crowded, cell-like conditions. The results show that macromolecular crowding affects protein-folding dynamics as well as overall protein shape. In crowded milieus, distinct conformational changes in VlsE are accompanied by secondary structure alterations that lead to exposure of a hidden antigenic region. Our work demonstrates the malleability of ''native'' proteins and implies that crowding-induced shape changes may be important for protein function and malfunction *in vivo*.

energy landscape theory | excluded volume effect | Lyme disease | macromolecular crowding | off-lattice model

The concentration of macromolecules inside cells is in the range of 80–400 mg/ml (1, 2), which corresponds to a volume occupancy of 5–40% (3). In such a crowded environment, the average spacing between macromolecules is much smaller than the size of the macromolecules themselves; therefore, any reactions that depend on available volume will be stimulated by macromolecular crowding effects (4, 5). Experimental and theoretical work has demonstrated strong effects of macromolecular crowding on thermodynamics and kinetics of many biological processes (1, 6–8). It has been established that the major result of macromolecular crowding is a stabilizing effect on the folded state of the protein due mostly to unfavorable effects on the unfolded-state ensemble (5, 9, 10). However, to identify the functional forms of proteins in cell-like conditions, the ability of macromolecular crowding to induce protein structural changes must be tested.

Combining *in vitro* and *in silico* methods, here we provide a description of how macromolecular crowding modulates structural changes in an aspherical protein. We propose the concept of protein-shape changes in cell-like environments as an amendment to the current framework of how macromolecular crowding affects protein properties. Using a protein model, we first present experimental results demonstrating that folding kinetics is enhanced in cell-like conditions *in vitro*, as predicted. Next, spectroscopic observations point to dramatic structural changes in the native protein at crowded conditions. Finally, to illustrate these phenomenological changes in the protein induced by macromolecular crowding with near all-atomistic detail, we implement molecular simulations and physics models.

Borrelia burgdorferi VlsE was selected as our model system for several reasons. First, it is an aspherical protein with marginal stability: It is best described as having an elongated football shape with a helical core surrounded by floppy loops at each end (11). Second, it unfolds reversibly *in vitro* in two-state (folded–unfolded) equilibrium and kinetic reactions (12, 13). Third, VlsE is proposed to be an important virulence factor upon mammalian infection. The *B. burgdorferi* spirochete is the causative bacteria of Lyme disease, a tick-born infection that is epidemic to regions of the U.S., Europe, and Asia. A specific diagnostic test for Lyme disease was derived from a 26-residue peptide region in VlsE named IR_6 (14). Surprisingly, for an immunodominant B cell epitope, this segment is hidden in the hydrophobic core of the VlsE crystal structure (11).

Results and Discussion

Low Levels of Crowding Affect Stability and Folding Speed of VlsE. Spectroscopic methods, fluorescence and far-UV CD, were used to monitor urea-induced unfolding of purified VlsE in the presence of Ficoll 70 (pH 7, 20°C) (Fig. 1*A*). Ficoll 70 is a common macromolecular crowding agent for *in vitro* studies. It is a highly cross-linked sucrose/epichlorohydrin copolymer that behaves like a semirigid sphere; it is chemically inert and not found to interact with proteins (15–20). The unfolding transitions detected by CD and fluorescence for VlsE are identical, and all reactions are reversible, corresponding to a two-state reaction. There is a shift in the transition midpoint to higher urea concentrations, and the unfolding-free energy increases in magnitude in the presence of Ficoll 70 (Table 1).

The mechanistic origin of the effects on stability was revealed via dynamic folding/unfolding experiments as a function of urea. All time-resolved reactions are best fit to single-exponential decay functions and provide rate constants (k_{obs}), the reactions are fully reversible, and there is no missing amplitude in the dead time. Semilogarithmic plots of ln k_{obs} versus [urea] are shown for 0 and 100 mg/ml Ficoll 70 in Fig. 1*B*. CD and fluorescence detection give identical results at each condition and all derived parameters agree with two-state behavior (Table 1). In the presence of 100 mg/ml Ficoll 70, the folding speed in water is 3-fold faster than without Ficoll 70, whereas there is no effect on the unfolding speed. A linear dependence of the logarithm of the folding rate constants as a function of Ficoll amount was found when folding was probed at 25 mg/ml increments between 0 and 100 mg/ml Ficoll 70 (Fig. 1*C*).

The *in vitro* VlsE kinetics are in excellent agreement with theoretical predictions based on the small WW domain peptide (7), where the folding kinetics for the WW domain was predicted to increase up to 3-fold after the inclusion of noninteracting spheres as crowding agents, yet unfolding speed was not affected (7). It is suggested that, because in a cell-like, colloidal solution with high-volume fraction (φ_c) of crowding agents, depletion-induced attractions (21) produce compact protein structures; thus, unfolded states are destabilized and folding rates are

Author contributions: M.S.C. and P.W.-S. designed research; D.H., M.P., and A.S. performed research; D.H., M.P., A.S., M.S.C., and P.W.-S. analyzed data; and M.S.C. and P.W.-S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

†D.H. and M.P. contributed equally to this work.

‖To whom correspondence should be sent at the present address: Department of Chemistry, Umeå University, 901 87 Umeå, Sweden. E-mail: pernilla.wittung@chem.umu.se.

This article contains supporting information online at www.pnas.org/cgi/content/full/0803672105/DCSupplemental.

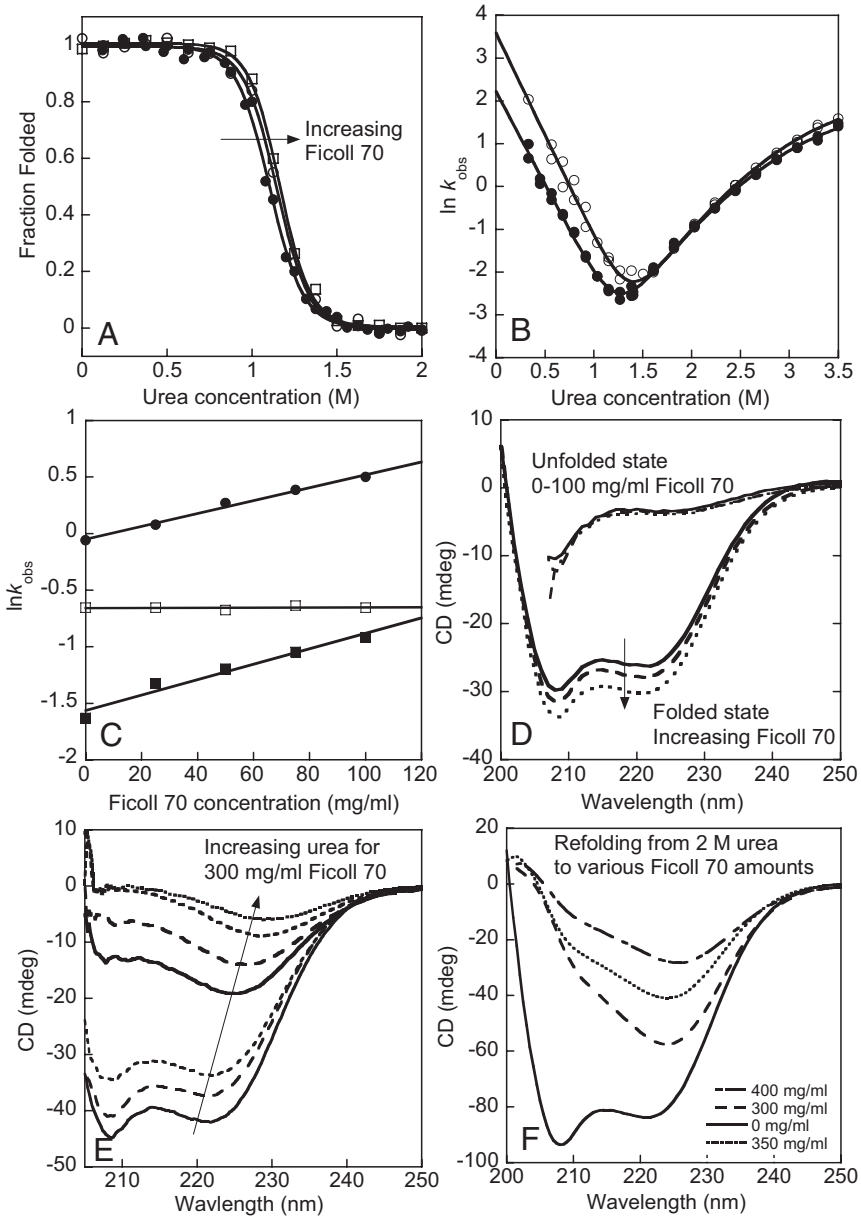

Fig. 1. *In vitro* effects of crowding on VlsE structure and folding. (*A*) Urea-induced unfolding of VlsE in 0 (filled circles), 50 (open squares), and 100 (open circles) mg/ml Ficoll 70. (*B*) Semilogarithmic plot of ln k_{obs} for VlsE unfolding/refolding as a function of urea in the presence of 0 (filled circles) and 100 (open circles) mg/ml Ficoll 70. (*C*) ln k_{obs} versus Ficoll 70 (25 mg/ml increments, 0–100 mg/ml) for denaturant jumps to three specific conditions: 0.6 M urea (filled circles), 2.2 M urea (open squares) and 0.9 M urea (filled squares). (*D*) Far-UV CD spectra of folded (in buffer) and unfolded (in 2 M urea) VlsE (pH 7, 20°C) in 0 (solid trace), 50 (dashed trace) and 100 (dotted trace) mg/ml Ficoll 70. (*E*) CD spectra of VlsE in 300 mg/ml Ficoll 70 as a function of urea (from 0 to 2.6 M urea). (*F*) Resulting CD spectra upon refolding attempts of VlsE unfolded in 2 M urea without Ficoll 70 into buffers containing various amounts of Ficoll 70 at >250 mg/ml (indicated in legend).

enhanced. Our *in vitro* data presented here on a much larger protein (341 residues versus 34) validates this prediction. The observation of similar transition-state placements with and without Ficoll 70 (Table 1, β_T values) supports that the folding path is not altered by this level of crowding agents.

Structural Content of Thermodynamic States of VlsE Depends on Crowding and Denaturant. There is an increase of helical secondary structure in the presence of increasing levels of Ficoll 70 in the absence of denaturant. This increase is illustrated in Fig. 1*D*, where the effects of up to 100 mg/ml Ficoll 70 on the far-UV CD signal of folded VlsE are shown. We reported earlier that the helical content in VlsE increases from 50% (as in crystal) to 80% upon inclusion of 400 mg/ml Ficoll 70 (22). In support of a nonspecific excluded volume effect, similar changes were found with dextran and PEG (22), but there was no increase in VlsE's secondary structure upon addition of high amounts of sucrose [i.e., the monomer unit of Ficoll (data not shown)] or glycerol (22).

A nonnative VlsE species with β-structure is populated at high levels of Ficoll 70 in the presence of urea but not at high levels of Ficoll 70 in the presence of buffer. Urea-induced unfolding experiments of VlsE in Ficoll 70 of 200 mg/ml or more show that the reactions are no longer folded-to-unfolded reactions; instead, the final state is better described as a "structured nonnative state" (Fig. 1*E*). The CD spectrum of this species has a negative peak near 225 nm, which reflects various β-structures (23). The same process takes place in thermal experiments with VlsE in the presence of ≥200 mg/ml Ficoll 70; based on CD

Table 1. Summary of *in vitro* effects of Ficoll 70 on VlsE folding

Condition	$[urea]_{1/2}$, M	$\Delta G_U(H_2O)$, kJ/mol	m_{eq}, kJ/mol·M	$k_F(H_2O)$, s^{-1}	$k_U(H_2O)$, s^{-1}	$m_F + m_U$, kJ/mol·M	β_T	ΔG_{KIN}, kJ/mol
Buffer	1.10 ± 0.04	24 ± 2	22 ± 1	7 ± 1	26×10^{-5}	22 ± 1	0.47	25 ± 2
Ficoll 70								
50 mg/ml	1.17 ± 0.03	26 ± 3	22 ± 3	—	—	—	—	—
100 mg/ml	1.20 ± 0.04	29 ± 1	24 ± 4	20 ± 3	20×10^{-5}	23 ± 1	0.49	29 ± 2

Equilibrium [$[urea]_{1/2}$, $\Delta G_U(H_2O)$, and m_{eq}] and kinetic [$k_F(H_2O)$, $k_U(H_2O)$, and $m_F + m_U$] parameters for VlsE urea-induced unfolding/refolding (pH 7, 20°C) and thermal unfolding (pH 7, no urea) as a function of crowding agent Ficoll 70. Parameters were obtained from two-state fits to the equilibrium and kinetic data. $\Delta G_{KIN}(H_2O)$ was calculated as $RT\ln[k_F(H_2O)/k_U(H_2O)]$, and Tanford β values (β_T) was calculated as $m_F/(m_F + m_U)$. —, not applicable.

spectra, the folded structure converts to the β-rich species at high temperatures (data not shown). Interestingly, when refolding of VlsE denatured in buffer with 2 M urea (no Ficoll 70) was triggered by dilutions into buffer solutions containing increasing amounts of Ficoll 70, the nonnative β-rich structures were detected for refolding conditions, including ≥250 mg/ml Ficoll 70 (Fig. 1*F*).

***In Silico* Analysis of Crowding-Induced Structural Changes in VlsE.** To explain the physics behind the macromolecular crowding effects that account for the *in vitro* observed structural changes in VlsE, we used a statistical physics approach of the energy landscape theory (24, 25) and aided by molecular simulations. To model VlsE from the crystal structure (11), we used a low-resolution (26, 27) protein representation in which all nonbonded interactions are considered and depend on sequence details. Ficoll 70 is modeled as a hard sphere (28) and provides nonspecific (hard-core repulsive) interactions in the molecular simulations [*Materials and Methods* and supporting information (SI) *Text*].

To analyze the structural properties of populated states in the simulated ensemble, the overlap function (χ), a microscopic measure of the similarity to the crystal structure ($\chi = 0$ for being similar; otherwise, $\chi = 1$) (11), and the radius of gyration (R_g) are used to characterize the energy landscape of VlsE (Fig. 2). We also use the shape (S) and asphericity (Δ) parameters to distinguish substantial changes in overall geometry of VlsE among those populated states ($S = \Delta = 0$ is a sphere, $\Delta > 0$ is aspherical, $S > 0$ is a prolate ellipsoid, and $S < 0$ is an oblate ellipsoid; see *Materials and Methods* for details). At low temperature ($k_BT/\varepsilon = 1$) in bulk, the shape of the dominant VlsE ensemble conformations ($0.00 < \chi < 0.10$ and $19.00 < R_g < 21.00$ Å; labeled "C" in Fig. 2*A*) resemble a football ($S = 0.34 \pm 0.00$; $\Delta = 0.31 \pm 0.00$). Second to the C state are ensemble conformations ($0.13 < \chi < 0.18$ and $18.00 < R_g < 21.00$ Å; labeled "B" in Fig. 2*A*) similar to a bent bean ($S = 0.14 \pm 0.00$; $\Delta = 0.17 \pm 0.00$). The ensemble conformations of the least populated state ($0.20 < \chi < 0.30$ and $18.00 < R_g < 19.30$ Å; labeled "X" in Fig. 2*A*) are spherical ($S = 0.02 \pm 0.00$; $\Delta = 0.02 \pm 0.00$) and most divergent from the crystal structure. At the highest temperature ($k_BT/\varepsilon = 1.47$), only unfolded states (labeled "U" in Fig. 2*C*) are populated. Analysis of the unfolded state ensembles at $k_BT/\varepsilon = 1.47$ in presence and absence of 25% crowders reveals, as expected, a shift in the distribution of R_g values toward more compact structures when crowding agents are present (Fig. S1).

Analysis of VlsE States Found on Computational Free Energy Landscape. Two-dimensional, free-energy diagrams for VlsE as a function of χ and R_g at three different temperatures (T) and ϕ_c of 0% (bulk), 15%, and 25% are shown in Fig. 2. At the lowest T ($k_BT/\varepsilon = 1$), φ_c stabilizes the C state. At a somewhat elevated T ($k_BT/\varepsilon = 1.13$ for bulk and 15%; $k_BT/\varepsilon = 1.2$ for 25% of crowders; a complete set of data for k_BT/ε of 1.13 and 1.2 is shown in Fig. S2), the population of the bean-like conformations starts to dominate over the football-like state. To reveal where in VlsE changes occurs in the bean structure, we computed the number of nonnative excess helical contacts, H_{nn} by using difference contact maps derived from the coarse-grained ensemble structures. H_{nn} is defined as the number of residues that form [nonnative] helical contacts in the bean structure not found in the crystal structure ensemble of VlsE. H_{nn} for the bean structure at $\varphi_c = 15\%$ and $k_BT/\varepsilon = 1.13$ is 32, which corresponds to an increase of helical structure of ≈30% (113 helical contacts in the C state and 145 helical contacts in the B state). In Fig. 3*A*, these 32 residues obtaining helical contacts in the *in silico* bean structure are indicated by using the crystal structure as the template. (For visualization purposes, a ribbon diagram of an arbitrarily chosen reconstructed all-atom structure of the B state with projections of the new helical contacts is provided in Fig. S3.) The new helical interactions in the B state are found in loop regions and in the end of helices. This analysis provides a structural prediction where increased helicity in VlsE is occurring *in vitro* and perhaps *in vivo* in low levels of crowding.

The appearance of excess nonnative helical contacts in the presence of crowding agents in both *in vitro* and *in silico* studies

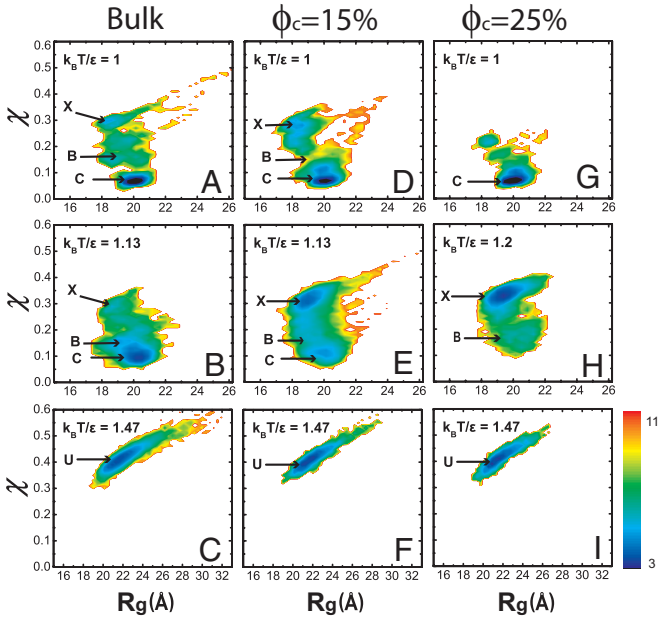

Fig. 2. Free-energy landscape for VlsE as a function of T and crowding level. Free-energy diagram as a function of the radius of gyration R_g and the overlap function (χ) for $\varphi_c = 0\%$ (bulk), 15%, 25% at various temperatures expressed in k_BT/ε. χ measures the deviation of structural similarity to the crystal (PDB ID code 1L8W); 0 means same as crystal. The color is scaled by k_BT where k_B is the Boltzmann's constant and ε is 0.6 kcal/mol. The native football-shaped species is labeled C (based on 1L8W with the omitted loop inserted), the bean structure is labeled B, the spherical state is named X, and the unfolded state is indicated by U. Note the different x axis scale in *C*, *F*, and *I* versus in *A*, *B*, *D*, *E*, *G*, and *H*.

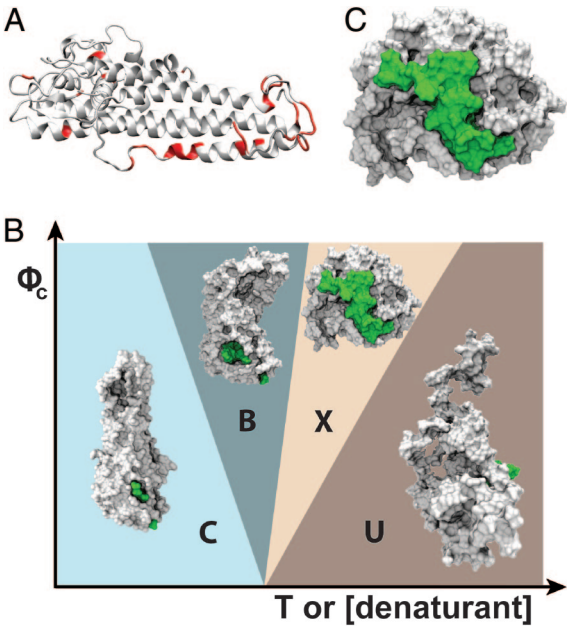

Fig. 3. Different nonnative structures induced by crowding and perturbations. (*A*) Nonnative excess helical contents, H_{nn}, are shown in red for the *in silico* bean structure at $\varphi_c = 15\%$ and $k_BT/\varepsilon = 1.13$, with the crystal structure as a structural template. (*B*) A schematic phase diagram of VlsE conformations in the φ_c–T (or urea) plane that agrees with both *in vitro* and *in silico* data. The solvent-accessible surface area of VlsE is in white and the antigenic IR_6 sequence is in green for all configurations. We use the same annotations as we did in Fig. 2 for the C, B, X, and U states. (*C*) Surface model of a high-resolution, all-atom reconstruction of the spherical state of VlsE, with the antigenic region IR_6, residues 272–299, highlighted in green. The surface exposure of IR_6 in this reconstructed structure is 41%.

was not the case in our recent report on macromolecular crowding effects on the native-state structural content of another protein, apoflavodoxin (28). Instead, in the presence of crowding agents, this protein became more compact and the structural content increased toward that found in the crystal structure. We argue that because apoflavodoxin is a spherical protein, no dramatic shape changes were observed in response to the isotropic forces generated by crowding. Because VlsE has an elongated shape, the depletion-induced attractions due to macromolecular crowding are able to promote more dramatic shape changes to reach the most favorable states.

When T is further increased ($k_BT/\varepsilon = 1.2$) and high φ_c is reached ($\varphi_c = 25\%$), the spherical X state becomes most populated (Fig. 2*H*). This species is formed by "breaking" the bean-like structure in the middle; thereby, both of the pointy ends of the protein are brought inwards, resulting in a compact spherical structure. Contact maps of the differences between the football (C state) and the bean (B state) and between the football and the spherical (X state) VlsE structures in terms of both native and nonnative interactions reveal that, whereas the B state gains nonnative helical content (as discussed above), the spherical X state has lost native helical interactions and instead has attained more nonnative interactions that appear to correspond to β-strand contacts (Fig. S4). A reconstructed high-resolution ribbon model of the X state is shown in Fig. S4, with the residues predicted to adopt β-strand contacts indicated. Upon comparing the helical and β-strand contents in the ensemble-averaged coarse-grained structures of the B and X states, we find that the helical content is reduced by a factor of $\approx$4 when going from B to X (from 145 to 36 predicted helical contacts), whereas 69 new β-strand contacts appear in the X state.

Correlation Between *in Vitro* and *in Silico* Observations. Remarkably, the *in vitro* behavior agrees with the computed free-energy landscape for VlsE; this concurrence may validate the underlying physics of the observed effects. We reasoned that T in the simulations and [urea] in the *in vitro* experiments represents similar factors that nonspecifically destabilize VlsE. In several cases, chemical and thermal perturbation mechanisms of a protein are similar (29, 30). Heating experiments demonstrated this similarity to be true for VlsE; at high temperature and in urea in the presence of large amounts of Ficoll 70, a nonnative structure forms that contains β-like structure. In Fig. 3*B*, we provide a phase diagram for VlsE conformations (i.e., football, C; bean, B; sphere, X; and unfolded, U) in φ_c–T (or urea) space that agrees with both computational and experimental results. Similar to a recent speculation on the complex behavior of helix formation in cell-like environments (31), the richness of our phase diagram is likely caused by the competition of φ_c and T (or [urea]). φ_c induces intramolecular attractions and the latter is inversely proportional to protein stability. Due to these competing factors, the critical region for structural changes becomes narrow.

Shape Changes in VlsE May Relate to Its Function. When VlsE is attached to the intact *B. burgdorferi* spirochete, the dominant antigenic region IR_6 is cryptic. Still, antibodies from infected serum samples are able to recognize IR_6 *in vivo* (14, 32). Although the IR_6 segment is hidden in the crystal structure (13% exposure to solvent), remarkably we find that this stretch is flanked out of the variable regions upon shape changes and becomes surface-exposed in the spherical X state of VlsE. Using the reconstructed high-resolution, all-atomistic model of X, we computed the accessible solvent area of the IR_6 region to be $31.7 \pm 1.3\%$ with standard deviation of 5.9% (Fig. 3*C*). Analysis of 20 randomly selected snapshots of the X state reveals that the IR_6 surface exposure range between 23% and 45% in these structures. Upon the release of VlsE into the host, the crowded milieu may trigger shape changes (mixture of B and X conformations) allowing for IR_6 exposure, which explains why lymphocyte receptors can access this region *in vivo* and trigger antibody generation in Lyme disease patients (14, 32).

Conclusions

The crowded environment in cells is predicted to affect folding and binding reactions of proteins due to excluded volume effects (4, 5, 33–36). Our discovery of high plasticity of an aspherical protein at crowded conditions *in vitro* and *in silico* implies that *in vivo* protein shape changes may easily occur depending on the environment. Shape changes are one way nature controls protein function, much like different facets can be exposed in an origami "cootie catcher." We speculate that it may be possible to develop "targeted crowding" as a precise tool to manipulate protein conformations *in situ* to turn on or off, as desired, specific activities and signaling cascades.

Materials and Methods

***In Vitro* Experiments.** VlsE was prepared as previously described (12). Far-UV CD (model J810 CD spectrometer; Jasco) was measured in a 1-mm cell, fluorescence (Eclipse fluorescence spectrophotometer; Varian) was measured in a 1-cm cell with excitation at 280 nm, and stopped-flow mixing experiments were executed (Pi-Star; Applied Photophysics) in far-UV CD and emission modes. Buffer was 20 mM phosphate, pH 7, at 20°C; final protein concentration was most often 10 μM. Urea concentrations were measured by refractive index. The equilibrium unfolding data as a function of Ficoll 70 was fit to two-state equations to reveal $\Delta G_U(H_2O)$ and m values. For kinetics, VlsE was first denatured in urea [$\pm$100 mg/ml Ficoll 70] and equilibrated for 16 h. For both refolding and unfolding, VlsE samples were mixed with solutions (with the same amount of Ficoll 70 as at the start) containing appropriate amounts of urea/buffer to trigger refolding or unfolding. VlsE heating experiments

were detected by far-UV CD in 0, 100, 200, 300, and 400 mg/ml Ficoll 70 using a scan rate of 1 deg/min.

***In Silico* Work.** A coarse-grained model of VlsE was constructed with a C_α side-chain model (SCM) (26, 27). Ficoll 70 is modeled as hard spheres that provide nonspecific repulsive interactions in the simulations. The thermodynamic properties of VlsE in bulk and in crowded environments (volume fractions, φ_c, of 0%, 15%, and 25%) were studied by molecular simulations in which the Langevin dynamics equation of motion at a low friction limit is used (37). Methods for preparation of equilibrated boxes of crowders and protein can be found in previous studies (7, 28). For an enhancement of sampling efficiency, the replica exchange method (38, 39) that simultaneously incorporates molecular simulations on 22 different temperatures ($0.83 < k_BT/\varepsilon < 1.53$) is used. Each replica generates ≈20,000 conformations for data to converge. Thermodynamic properties were analyzed by the weighted histogram analysis method (40, 41).

Coarse-Grained Model of VlsE. The sequence between residues 92 and 113 that defines the loop region of a VlsE [Protein Data Bank (PDB) ID code 1L8W] lacks structure, and the coordinates are omitted from the PDB. This loop region was inserted to the protein and energetically minimized under a CHARMM force field (42, 43) with the Insight II package (Accelrys). This structure was used to generate the coarse-grained model of VlsE. In this model, one bead is taken from the C_α position and the other is justified by the center of mass of a side chain. The potential energy of this system is $E_p + E_{cc} + E_{pc}$, where E_p, E_{cc}, and E_{pc} are the potential energies of the protein, interactions between crowding agents, and interactions between the protein and crowding agents, respectively. The potential energy of the protein is $E_p = E_s + E_{nonbonded}$. The structural energy, E_s, is the sum of the bond-length potential, the bond-angle potential, the dihedral potential, and the chiral potential (see *SI Text* for details) (26, 44), and the equilibrium values for each term are taken from the native structure of VlsE. As for the inserted loop, we kept all of the terms in E_s except the dihedral potential to capture the physics of a flexible loop. The nonbonded term, $E_{nonbonded}$, includes the side chain–side chain interaction (E_{ss}), the side chain–backbone interaction (E_{sb}), and the backbone–backbone interaction (E_{bb}). Solvent-mediated interaction between pairs of side chains, ε_{ij}, is based on the Betancourt–Thirumalai statistical potential (45). This statistical potential addresses sequence variations where the reference interaction, $\varepsilon = 0.6$ kcal/mol, is based on the Thr–Thr pairwise interactions. E_{ss} of two interacting side chains, *i* and *j*, that include both native and nonnative pairs at a distance *r* follows

$$E_{ss}^{ij} = \varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - 2\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right], \qquad [1]$$

where $\sigma_{ij} = f(\sigma_i + \sigma_j)$ and σ_i and σ_j are the van der Waals (vdW) radii of the side chains. To avoid clashes between bulky side chains, $f = 0.9$.

E_{bb} is the multiplication of a Lennard-Jones term similar to Eq. **1** and an angular term, $A(\rho)$, that assesses major secondary structures (see *SI Text* for details). ρ is the pseudodihedral angle that defines the alignment of a pair of interacting backbone hydrogen bonds by four C_α beads (26). Repulsive interactions are used for E_{sb}, E_{cc}, and E_{pc}. The repulsive potential between two particles *i* and *j* at a distance r_{ij} follows

$$E^{ij} = \varepsilon\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}, \qquad [2]$$

where $\sigma_{ij} = \sigma_i + \sigma_j$; σ_i and σ_j are vdW radii of interacting beads. $\varepsilon = 0.6$ kcal/mol.

Asphericity and Shape Parameters. To determine the shape of a given state, we use two rotationally invariant quantities: shape parameter, *S*, and asphericity parameter, Δ. These quantities are calculated with the use of the inertia tensor (46):

$$T_{\alpha\beta} = \frac{1}{2N^2}\cdot\sum_{i,j=1}^{N}(r_{i\alpha} - r_{j\alpha})\cdot(r_{i\beta} - r_{j\beta}), \qquad [3]$$

where *N* is the number of beads in the protein, $r_{i\alpha}$ is the αth position of bead *i* and α, β are the coordinates *x*, *y*, *z*. The eigenvalues of *T* and λ_i are the squares of the three principal radii of gyration. Therefore,

$$R_g^2 = \mathrm{tr}(T) = \sum_{i=1}^{3}\lambda_i. \qquad [4]$$

Asphericity Δ is calculated by using

$$\Delta = \frac{3\sum_{i=1}^{3}(\lambda_i - \bar{\lambda})^2}{2(\mathrm{tr}T)^2}, \qquad [5]$$

where $\bar{\lambda} = \mathrm{tr}(T)/3$. The shape parameter *S* is calculated by using

$$S = 27\cdot\frac{\prod_{i=1}^{3}(\lambda_i - \bar{\lambda})}{(\mathrm{tr}(T))^3}. \qquad [6]$$

The parameters *S* and Δ lie in the following ranges: $0 \le \Delta \le 1$ and $-0.25 \le S \le 2$. For a perfect sphere we have $S = \Delta = 0$. Deviation of Δ from 0 is an indication of the extent of anisotropy for the state under study. Negative values of *S* refer to oblate ellipsoids, whereas positive values correspond to prolate ellipsoids.

Reconstruction of All-Atom Protein Model. The SCAAL (side chain-C_α to all-atom) method was used to reconstruct all-atom protein structures from coarse-grained simulation snapshots. This method implements a constrained energy minimization algorithm on the C_α and one heavy atom in the side chain (the one in closest proximity to the center of mass of the side chain). Starting from a known all-atom protein structure, constraints are applied on two heavy atoms per residue (except glycine). These two heavy atoms are constrained to the coordinates of the C_α and side-chain beads in a selected coarse-grained ensemble structure. The spring constant for the constraint dynamics is 100 kcal/mol·Å^2. The minimization was performed by using the conjugate gradient method from NAMD (4) using the CHARMM (42, 43) force field with until the tolerance reaches 10^{-8}. To test the accuracy of this reconstruction method, the crystal structure 1L8W was first coarse-grained into C_α–SCM and then reconstructed back to all-atom representation. The root-mean-square-deviation of the reconstructed all-atom representation with respect to the original crystal structure, based on all-heavy atoms, is 0.6 Å, which is much smaller than the resolution of the crystal structure solved by x-ray diffraction (2.3 Å). Surface accessible area was computed by using GETAREA1.1 (48).

ACKNOWLEDGMENTS. The computations were performed on the National Science Foundation Terascale Computing System (Teragrid; TG-MCB070066N) and Texas Learning Computation Center (TLC^2) at the University of Houston. This work was supported by Robert A. Welch Foundation Grant C-1588 (to P.W.S.). M.S.C. was supported by the University of Houston and a Texas Center for Superconductivity seed grant.

1. van den Berg B, Ellis RJ, Dobson CM (1999) Effects of macromolecular crowding on protein folding and aggregation. *EMBO J* 18:6927–6933.
2. Rivas G, Ferrone F, Herzfeld J (2004) Life in a crowded world. *EMBO Rep* 5:23–27.
3. Ellis RJ, Minton AP (2003) Cell biology: Join the crowd. *Nature* 425:27–28.
4. Zimmerman SB, Minton AP (1993) Macromolecular crowding: Biochemical, biophysical, and physiological consequences. *Annu Rev Biophys Biomol Struct* 22:27–65.
5. Minton AP (2005) Models for excluded volume interaction between an unfolded protein and rigid macromolecular cosolutes: Macromolecular crowding and protein stability revisited. *Biophys J* 88:971–985.
6. Uversky VN, E MC, Bower KS, Li J, Fink AL (2002) Accelerated alpha-synuclein fibrillation in crowded milieu. *FEBS Lett* 515:99–103.
7. Cheung MS, Klimov D, Thirumalai D (2005) Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proc Natl Acad Sci USA* 102:4753–4758.
8. Bokvist M, Grobner G (2007) Misfolding of amyloidogenic proteins at membrane surfaces: The impact of macromolecular crowding. *J Am Chem Soc* 129:14848–14849.
9. Hall D, Minton AP (2003) Macromolecular crowding: qualitative and semiquantitative successes, quantitative challenges. *Biochim Biophys Acta* 1649:127–139.
10. Zhou HX, Dill KA (2001) Stabilization of proteins in confined spaces. *Biochemistry* 40:11289–11293.
11. Eicken C, *et al*. (2002) Crystal structure of Lyme disease variable surface antigen VlsE of Borrelia burgdorferi. *J Biol Chem* 28:28.
12. Jones K, Wittung-Stafshede P (2003) The largest protein observed to fold by two-state kinetic mechanism does not obey contact-order correlation. *J Am Chem Soc* 125:9606–9607.
13. Jones K, Guidry J, Wittung-Stafshede P (2001) Characterization of surface antigen from Lyme disease spirochete Borrelia burgdorferi. *Biochem Biophys Res Commun* 289:389–394.

14. Liang FT, *et al*. (1999) Sensitive and specific serodiagnosis of Lyme disease by enzyme-linked immunosorbent assay with a peptide based on an immunodominant conserved region of Borrelia burgdorferi vlsE. *J Clin Microbiol* 37:3990–3996.
15. Venturoli D, Rippe B (2005) Ficoll and dextran vs. globular proteins as probes for testing glomerular permselectivity: Effects of molecular size, shape, charge, and deformability. *Am J Physiol* 288:F605–F613.
16. Luby-Phelps K, Castle PE, Taylor DL, Lanni F (1987) Hindered diffusion of inert tracer particles in the cytoplasm of mouse 3T3 cells. *Proc Natl Acad Sci USA* 84:4910–4913.
17. Monterroso B, Minton AP (2007) Effect of high concentration of inert cosolutes on the refolding of an enzyme: Carbonic anhydrase B in sucrose and Ficoll 70. *J Biol Chem* 282:33452–33458.
18. van den Berg B, Wain R, Dobson CM, Ellis RJ (2000) Macromolecular crowding perturbs protein refolding kinetics: Implications for folding inside the cell. *EMBO J* 19:3870–3875.
19. Zhou BR, Liang Y, Du F, Zhou Z, Chen J (2004) Mixed macromolecular crowding accelerates the oxidative refolding of reduced, denatured lysozyme: Implications for protein folding in intracellular environments. *J Biol Chem* 279:55109–55116.
20. Du F, *et al*. (2006) Mixed macromolecular crowding accelerates the refolding of rabbit muscle creatine kinase: Implications for protein folding in physiological environments. *J Mol Biol* 364:469–482.
21. Asakura S, Oosawa F (1954) On interaction between two bodies immersed in a solution of macromolecules. *J Chem Phys* 22:1255–1256.
22. Perham M, Stagg L, Wittung-Stafshede P (2007) Macromolecular crowding increases structural content of folded proteins. *FEBS Lett* 581:5065–5069.
23. Manning MC, Illangasekare M, Woody RW (1988) Circular dichroism studies of distorted alpha-helices, twisted beta-sheets, and beta turns. *Biophys Chem* 31:77–86.
24. Onuchic JN, Luthey-Schulten Z, Wolynes PG (1997) Theory of protein folding: The energy landscape perspective. *Annu Rev Phys Chem* 48:545–600.
25. Brooks CL, III, Gruebele M, Onuchic JN, Wolynes PG (1998) Chemical physics of protein folding. *Proc Natl Acad Sci USA* 95:11037–11038.
26. Cheung MS, Finke JM, Callahan B, Onuchic JN (2003) Exploring the interplay of topology and secondary structural formation in the protein folding problem. *J Phys Chem B* 107:11193–11200.
27. Klimov DK, Newfield D, Thirumalai D (2002) Simulations of beta-hairpin folding confined to spherical pores using distributed computing. *Proc Natl Acad Sci USA* 99:8019–8024.
28. Stagg L, Zhang S-Q, Cheung MS, Wittung-Stafshede P (2007) Molecular crowding enhances native structure and stability of alpha/beta protein flavodoxin. *Proc Natl Acad Sci USA* 104:18976–18981.
29. Perham M, Chen M, Ma J, Wittung-Stafshede P (2005) Unfolding of heptameric co-chaperonin protein follows ''fly casting'' mechanism: Observation of transient nonnative heptamer. *J Am Chem Soc* 127:16402–16403.
30. Zong C, Wilson CJ, Shen T, Wolynes PG, Wittung-Stafshede P (2006) Phi-value analysis of apo-azurin folding: Comparison between experiment and theory. *Biochemistry* 45:6458–6466.
31. Snir Y, Kamien RD (2005) Entropically driven helix formation. *Science* 307:1067.
32. Philipp MT, *et al*. (2001) Antibody response to IR6, a conserved immunodominant region of the VlsE lipoprotein, wanes rapidly after antibiotic treatment of Borrelia burgdorferi infection in experimental animals and in humans. *J Infect Dis* 184:870–878.
33. Sasahara K, McPhie P, Minton AP (2003) Effect of dextran on protein stability and conformation attributed to macromolecular crowding. *J Mol Biol* 326:1227–1237.
34. Spencer DS, Xu K, Logan TM, Zhou HX (2005) Effects of pH, salt, and macromolecular crowding on the stability of FK506-binding protein: An integrated experimental and theoretical study. *J Mol Biol* 351:219–232.
35. Qu Y, Bolen D (2002) Efficacy of macromolecular crowding in forcing proteins to fold. *Biophys Chem* 101–102:155–165.
36. Tokuriki N, *et al*. (2004) Protein folding by the effects of macromolecular crowding. *Protein Sci* 13:125–133.
37. Veitshans T, Klimov D, Thirumalai D (1997) Protein folding kinetics: Timescales, pathways and energy landscapes in terms of sequence-dependent properties. *Fold Des* 2:1–22.
38. Sanbonmatsu KY, Garcia AE (2002) Structure of Met-Enkephalin in explicit aqueous solution using replica exchange molecular dynamics. *Proteins* 46:225–234.
39. Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics methods for protein folding. *Chem Phys Lett* 314:141–151.
40. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM (1992) The weighted histogram analysis method for free-energy calculations on biomolecules I. The method. *J Comput Chem* 13:1011–1021.
41. Chodera JD, Swope WC, Pitera JW, Seok C, Dill KA (2007) Use of the weighted histogram analysis method for the analysis of simulated and parallel tempering simulations. *J Chem Theory Comput* 3:26–41.
42. Brooks BR, *et al*. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 4:187–217.
43. MacKerel AD, Jr, *et al*. (1998) CHARMM: The energy function and its parameterization with an overview of the program. *The Encyclopedia of Computational Chemistry*, ed von Ragué Schleyer P (Wiley, Chichester, NY), Vol 1, pp 271–277.
44. Cheung MS, Thirumalai D (2006) Nanopore-protein interactions dramatically alter stability and yield of the native state in restricted spaces. *J Mol Biol* 357:632–643.
45. Betancourt MR, Thirumalai D (1999) Exploring the kinetic requirements for enhancement of protein folding rates in the GroEL cavity. *J Mol Biol* 287:627–644.
46. Dima RI, Thirumalai D (2004) Asymmetry in the shapes of folded and denatured states of proteins. *J Phys Chem B* 108:6564–6570.
47. Phillips JC, *et al*. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802.
48. Fraczkiewicz R, Braun W (1998) Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J Comput Chem* 19:319–333.